Express Mail Label No. **EL339142538US**

UTILITY PATENT APPLICATION TRANSMITTAL
(Large Entity)

(Only for new nonprovisional applications under 37 CFR 1.53(b))

Docket No. YO998-525
Total Pages in this Submission

TO THE ASSISTANT COMMISSIONER FOR PATENTS
Box Patent Application
Washington, D.C. 20231

Transmitted herewith for filing under 35 U.S.C. 111(a) and 37 C.F.R. 1.53(b) is a new utility patent application for an invention entitled:

MESSAGE LOGGING FOR RELIABLE MULTICASTING ACROSS A ROUTING NETWORK

and invented by:

Guruduth Somasekhara Banavar, Tushar Deepak Chandra, Kevan Lee Miller, Robert Evan Strom, Daniel Charles Sturman and Michael James Ward

If a CONTINUATION APPLICATION, *check appropriate box and supply the requisite information:*

☐ **Continuation** ☐ **Divisional** ☐ **Continuation-in-part (CIP)** of prior application No.: _____

Which is a:

☐ **Continuation** ☐ **Divisional** ☐ **Continuation-in-part (CIP)** of prior application No.: _____

Which is a:

☐ **Continuation** ☐ **Divisional** ☐ **Continuation-in-part (CIP)** of prior application No.: _____

Enclosed are:

Application Elements

1. ☒ Filing fee as calculated and transmitted as described below

2. ☒ Specification having 47 pages and including the following:

   a. ☒ Descriptive Title of the Invention
   b. ☒ Cross References to Related Applications *(if applicable)*
   c. ☐ Statement Regarding Federally-sponsored Research/Development *(if applicable)*
   d. ☐ Reference to Microfiche Appendix *(if applicable)*
   e. ☒ Background of the Invention
   f. ☒ Brief Summary of the Invention
   g. ☒ Brief Description of the Drawings *(if drawings filed)*
   h. ☒ Detailed Description
   i. ☒ Claim(s) as Classified Below
   j. ☒ Abstract of the Disclosure

UTILITY PATENT APPLICATION TRANSMITTAL
(Large Entity)

(Only for new nonprovisional applications under 37 CFR 1.53(b))

Docket No. YO998-525
Total Pages in this Submission

Application Elements (Continued)

3. ☒ Drawing(s) *(when necessary as prescribed by 35 USC 113)*
   a. ☒ Formal Number of Sheets **Ten (10)**
   b. ☐ Informal Number of Sheets _____
4. ☒ Oath or Declaration
   a. ☒ Newly executed *(original or copy)* ☐ Unexecuted
   b. ☐ Copy from a prior application (37 CFR 1.63(d)) *(for continuation/divisional application only)*
   c. ☒ With Power of Attorney ☐ Without Power of Attorney
   d. ☐ *DELETION OF INVENTOR(S)*
   Signed statement attached deleting inventor(s) named in the prior application, see 37 C.F.R. 1.63(d)(2) and 1.33(b).
5. ☐ Incorporation By Reference *(usable if Box 4b is checked)*
   The entire disclosure of the prior application, from which a copy of the oath or declaration is supplied under Box 4b, is considered as being part of the disclosure of the accompanying application and is hereby incorporated by reference therein.
6. ☐ Computer Program in Microfiche *(Appendix)*
7. ☐ Nucleotide and/or Amino Acid Sequence Submission *(if applicable, all must be included)*
   a. ☐ Paper Copy
   b. ☐ Computer Readable Copy *(identical to computer copy)*
   c. ☐ Statement Verifying Identical Paper and Computer Readable Copy

Accompanying Application Parts

8. ☒ Assignment Papers *(cover sheet & document(s))*
9. ☐ 37 CFR 3.73(B) Statement *(when there is an assignee)*
10. ☐ English Translation Document *(if applicable)*
11. ☒ Information Disclosure Statement/PTO-1449 ☒ Copies of IDS Citations
12. ☐ Preliminary Amendment
13. ☒ Acknowledgment postcard
14. ☒ Certificate of Mailing
    ☐ First Class ☒ Express Mail *(Specify Label No.):* **EL339142538US**

UTILITY PATENT APPLICATION TRANSMITTAL
(Large Entity)

(Only for new nonprovisional applications under 37 CFR 1.53(b))

Docket No. YO998-525
Total Pages in this Submission

Accompanying Application Parts (Continued)

15. ☐ Certified Copy of Priority Document(s) *(if foreign priority is claimed)*

16. ☒ Additional Enclosures *(please identify below):*

Associate Power of Attorney (2 pp.)

Fee Calculation and Transmittal

CLAIMS AS FILED

For	#Filed	#Allowed	#Extra	Rate	Fee
Total Claims	50	- 20 =	30	x $18.00	$540.00
Indep. Claims	8	- 3 =	5	x $78.00	$390.00
Multiple Dependent Claims (check if applicable) ☐					$0.00
				BASIC FEE	$760.00
OTHER FEE *(specify purpose)* _____					$0.00
				TOTAL FILING FEE	$1,690.00

☒ A check in the amount of **$1,690.00** to cover the filing fee is enclosed.

☒ The Commissioner is hereby authorized to charge and credit Deposit Account No. **08-1935** as described below. A duplicate copy of this sheet is enclosed.

- ☐ Charge the amount of _____ as filing fee.
- ☒ Credit any overpayment.
- ☒ Charge any additional filing fees required under 37 C.F.R. 1.16 and 1.17.
- ☐ Charge the issue fee set in 37 C.F.R. 1.18 at the mailing of the Notice of Allowance, pursuant to 37 C.F.R. 1.311(b).

Kevin P. Radigan
Signature

Dated: **March 30, 1999**

Kevin P. Radigan, Esq.
Reg. No. 31,789
HESLIN & ROTHENBERG, P.C.
5Columbia Circle
Albany, NY 12203
Tel.: (518) 452-5600
Fax: (518) 452-5579

cc:

MESSAGE LOGGING FOR RELIABLE MULTICASTING ACROSS A ROUTING NETWORK

Cross-Reference to Related Applications

This application contains subject matter which is related to the subject matter of the following United States patent applications, which are assigned to the same assignee of this application. Each of the below-listed applications is hereby incorporated herein by reference:

"Routing Messages Within A Network Using The Data Content Of The Message," by Chandra et al., filed November 20, 1997, serial no. 08/975,303;

"Message Sequencing For Ordered Multicasting Of A Message Across A Routing Network," by Banavar et al., co-filed herewith, serial no. , (attorney docket no. YO998-526);

"Quiescent Reconfiguration Of A Routing Network," by Miller et al., co-filed herewith, serial no. , (attorney docket no. YO998-527); and

"Non-Disruptive Reconfiguration Of A Publish/Subscribe System," by Miller et al., co-filed herewith, serial no. , (attorney docket no. YO999-124).

Technical Field

This invention relates in general to multicasting messages within a network and, in particular, to logging messages to persistent storage within a routing network to facilitate reliable delivery thereof to one or more clients. Further, the invention relates to reliably multicasting a message within a content-based routing network irrespective of destination information that may be included within the message.

Background of the Invention

Many network environments enable messages to be forwarded from one site within the network to one or more other sites using a multicast protocol. Typical multicast protocols send messages from one site to one or more other sites based on information stored within a message header. That is, each message has two components: the message header, which includes the routing information, including destination addresses or a predefined group name that is associated with a fixed list of destinations, and a data content, which is the data of the message. The routing information is read from the message header and is used to send the data content of the message to the specified destinations.

One example of a system that conventionally includes such a network environment is a publish/subscribe system. In publish/subscribe systems, publishers post messages and subscribers independently specify categories of events in which they are interested. The system takes the posted

messages and includes in each message header the destination information of those subscribers indicating interest in the particular message. The system then uses the destination information in the message to forward the message through the network to the appropriate subscribers.

In large systems, there may be many subscribers interested in a particular message. Thus, a large list of destinations needs to be added to the message header and used in forwarding the message. The use of the list, which can be even longer than the message itself, can degrade system performance. Other approaches have included the use of multicast groups, in which destinations are bound to a group name, and then that name is included in the message header. The message is then sent to all those destinations bound to the name. This technique has the disadvantage of requiring static groups of destinations, which restricts flexibility in many publish/subscribe systems.

Disclosure of the Invention

A publish/subscribe system consists of a network of message routers (or simply routers) connected via links in an arbitrary graph topology. A number of clients connect to the periphery of this router network and either publish or subscribe to messages. A message consists of a number of attributes, which are name-value pairs. A subscription specifies a predicate on the attributes of messages.

The router network is responsible for routing messages from publishers to interested subscribers based on matching events to subscription predicates. One embodiment of this

routing protocol, referred to as content-based routing, is described in the above-incorporated co-pending patent application entitled "Routing Messages Within A Network Using The Data Content Of The Message." Content-based routing proceeds as follows. From each router node at which a publisher is present, the system computes and stores a spanning tree to reach every other node in the network. All published messages from a particular publisher then follow the paths in the corresponding spanning tree, with each router node performing enough matching to determine which of its child routers should receive the published message. All paths preserve a first-in first-out (FIFO) ordering on messages from any one publisher to all subscribers.

Briefly summarized, the present invention is directed in one aspect to providing a mechanism for multicasting messages in a manner which is resilient to failures in the network, routers, or clients (e.g., publishers/subscribers). Resiliency is achieved by logging messages to persistent storage within the network prior to delivery thereof to one or more clients.

More particularly, provided herein is a method for routing messages within a network. The method includes: receiving a message; and routing the message to one or more clients of the network, the routing being based on data content of the message irrespective of any destination information that may be within the message, and being resilient to router or link failure within the network.

In another aspect, the invention comprises a method for routing messages within a network which includes: receiving

a message; logging the message to persistent storage within the network; and routing the message to one or more clients of the network after logging the message to persistent storage, wherein the logging is used to ensure resiliency in routing the message to one or more clients of the network notwithstanding failure at a router or link within the network.

In yet another aspect, a system of routing messages within a network is provided. The system includes means for receiving a message, and means for routing the message to one or more clients of the network. The routing is based on data content of the message irrespective of any destination information that may be within the message, and the means for routing is resilient to router or link failure within the network.

In still another aspect, a system of routing messages within a routing network is provided which includes means for receiving a message, means for logging the message to persistent storage within the routing network, and means for delivering the message to one or more clients of the network after logging thereof to the persistent storage. By logging the message to persistent storage prior to delivery thereof, a technique is provided which ensures resiliency to the routing of the message to the one or more clients notwithstanding router or link failure within the network.

In a further aspect, an article of manufacture is provided which includes at least one computer usable medium having computer readable program code means embodied therein for effecting routing of messages within the network. The

computer readable program code means in the article of manufacture includes: computer readable program code means for causing a computer to effect receiving a message; and computer readable program code means for causing a computer to effect routing the message to one or more clients of the network, the routing being based on data content of the message irrespective of any destination information that may be within the message, and being resilient to router or link failure within the network.

In a still further aspect, an article of manufacture is provided which includes at least one computer usable medium having computer readable program code means embodied therein for effecting routing of messages within a routing network. The computer readable program code means in the article of manufacture includes: computer readable program code means for causing a computer to effect receiving a message; computer readable program code means for causing a computer to effect logging the message to persistent storage within the routing network; and computer readable program code means for causing a computer to effect delivering the message to one or more clients of the network after the logging thereof, wherein the logging to persistent storage prior to delivery of the message provides resiliency to the routing network notwithstanding router or link failure within the network.

To restate, provided herein is a technique for logging messages within a routing network itself to ensure reliable multicasting across the network. Although principally described herein in connection with data-content routing of messages, the logging technique can be applied to other

routing systems, e.g., subject-based routing. The technique is scalable through the provision of multiple loggers within the network. Additionally, multiple loggers can be used to balance the load on the system. A system implementing the logging technique of the present invention allows senders and receivers to connect anywhere in the network regardless of the location of the loggers. Messages are preferably logged only once within a router network. Again, logging of messages occurs within the network itself between the sender and the one or more clients to which the message is directed.

The logging technique described herein allows a sender or receiver to choose the level of reliability (i.e., quality of service) required by either the sender or the clients. This choice allows a tradeoff of system resources versus quality of service. For a sender or receiver requesting "uniform delivery", the routing network either delivers the message to all clients or delivers the message to none of the clients. For a sender or receiver requesting "at most once delivery", the routing network detects and eliminates duplicate messages. "At least once delivery" quality of service is also provided by ensuring that loss of a message within the routing network is identified, and the sender is notified to retransmit the message. For a sender or receiver requesting "exactly once delivery" the routing network neither loses the message without notifying the sender nor makes a duplicate delivery of the message to a client. Again, the particular type of delivery can be selected by the sender or the receiving client.

Brief Description of the Drawings

The above-described objects, advantages and features of the present invention, as well as others, will be more readily understood from the following detailed description of certain preferred embodiments of the invention, when considered in conjunction with the accompanying drawings in which:

Fig. 1 depicts one example of a distributed network incorporating and using the resilient routing capability of the present invention;

Fig. 2 depicts one example of a spanning tree used in accordance with the principles of the present invention;

Fig. 3 depicts one embodiment of a distributed router network configured with the resilient routing capability of the present invention;

Fig. 4 depicts one embodiment of a router for use in accordance with the principles of the present invention in a router network such as depicted in **Fig. 3**;

Fig. 5 depicts one example of a router configured as a logger with message logging capability in accordance with the principles of the present invention;

Fig. 6a depicts one example of a client message buffer used by a router having clients connected thereto in accordance with the principles of the present invention;

Fig. 6b depicts one example of a logging acknowledgment (LACK) message sent by a logger to a router in accordance with the principles of the present invention;

Fig. 7 is a flowchart of one embodiment of a resilient message routing process in accordance with the principles of the present invention;

Fig. 8 is a flowchart of one embodiment of logging acknowledgment (LACK) processing in accordance with the principles of the present invention;

Fig. 9 is a flowchart of one example of processing for failure recovery within the routing network in accordance with the principles of the present invention;

Fig. 10 is a flowchart of an at most once delivery quality of service in accordance with the principles of the present invention; and

Fig. 11 is a flowchart of an at least once delivery quality of service in accordance with the principles of the present invention.

Best Mode for Carrying Out the Invention

As noted above, the present invention is in one aspect an extension of the routing capability disclosed in the above-incorporated application wherein messages are routed in a network based on the content of the data within the message. In particular, the message does not need to include destination information, such as destination

addresses or a group destination name. Instead, data within the message is used to traverse a data structure to determine the link or links over which the message is to be forwarded in order to reach the consumers (subscribers or clients) interested in the message.

In accordance with the principles of the present invention, this content-based routing capability is enhanced for multicasting messages in a manner which is resilient to failures in the network, its routers, or clients. The invention is referred to herein as comprising a "reliable routing capability" for a routing network such as the above-described content-based routing network. However, those skilled in the art will recognize that the logging technique presented hereinbelow for ensuring message delivery notwithstanding a failure within the network is equally applicable to other routing systems, such as a subject-based routing system. This reliable routing capability includes support for four "qualities of service": (1) "uniform delivery" under which the routing network either delivers the message to all receivers or none of the receivers, (2) "at most once delivery" under which the routing network detects and eliminates duplicate messages, (3) "at least once delivery" under which the routing network detects the loss of a message and notifies the sender to retransmit the message, and (4) "exactly once delivery" under which the routing network neither loses a message without notifying the sender nor makes duplicate deliveries to receivers.

One example of a distributed network 100 incorporating and using the reliable routing capability of the present invention is depicted in **Fig. 1** and described in detail

below. Network 100 includes, for instance, a plurality of computing units 102 coupled to one another by links 104.

Each link couples two computing units in the network, and each computing unit may have any number of links connected to it. Each link is bidirectional, i.e., a computing unit may send and receive messages on the link. Each link is also connection oriented, i.e., the computer units at the ends of the link are notified when the link fails. Furthermore, each link propagates messages in a first-in first-out manner. Each computing unit in the network is either a client computer (represented by the smaller ovals, such as those having addresses 101a, 101b), meaning that it has requested to receive messages whose content satisfies certain properties; or it is a router computer (represented by the larger ovals, such as 108a, 108b), meaning that it forwards messages received on one network link onto other links on the way to the client computer(s). The clients are collectively referred to herein as clients 101 and the routers are collectively referred to as routers 108. (For purposes of this discussion, if a single computing unit serves both as a router and as a client, these two separate functions will be considered as two computing units connected by a link.)

Each computing unit can be any type of computing unit that can be included in a network. For example, it can be an RS/6000 computing node or any other of various types of computing nodes, processors, computers or systems. The network can also include different types of computing units coupled to one another via the links. The links include,

for instance, TCP connections over IP links, as only one example.

Distributed network 100 can be included in various systems that require the passing of messages or data. These systems include, for instance, the internet or a content-based publish/subscribe system.

Content-based publish/subscribe systems improve the degree of decoupling between publishers and subscribers. In content-based publish/subscribe systems, subscriptions are specified in terms of predicates on the posted data, rather than in terms of subject identifiers supplied by the publisher. One example of a content-based publish/subscribe system is described in co-pending U.S. Patent Application Serial No. 08/975,280, entitled "Method And System For Matching Consumers To Events," Astley et al., which is hereby incorporated herein by reference in its entirety. The examples described herein are with reference to a content-based subscription system. However, these are only examples. The present invention can be employed with other types of systems without departing from the spirit of the present invention.

In one embodiment of the invention, each router 108 of network 100 (**Fig. 1**) has associated therewith a spanning tree, which lays out the best path (according to some criterion, such as latency) from the router to each client 101. In this embodiment, it is assumed that routers agree on a common criterion for measuring distance between nodes in the network. There may in fact be multiple spanning trees. For example, alternative spanning trees may specify

either backup routes, or peak load routes. Herein, it is assumed that one spanning tree is in effect for the routing of any particular message.

One example of a spanning tree, which is associated with router 108a (**Fig. 1**), is depicted in **Fig. 2**. As shown in **Fig. 2**, there is a path from router 102a to every other node in the spanning tree. A message to be routed from router 102a to one or more of the other nodes is routed via one or more of the links associated with router 102a, i.e., links 1-3. For example, if a message is to be routed from node 102a to node 101c, then link 2 is used. As a further example, if a message is to be forwarded toward client 101a, then link 3 is used.

One embodiment for building a spanning tree from a network topology, that is an arbitrary graph, is described in detail in Introduction to Algorithms, by Cormen, Leiserson, Rivert, Chapter 24, pp 498-513, Published by MIT Press (1990), which is hereby incorporated herein by reference in its entirety.

In addition to the spanning tree associated with each router, each router has a routing table. The routing table includes an entry for each client computer in the network. Each entry of the routing table associates a client address with the identifier of the network link constituting the next segment on the path in the spanning tree from the router to the client. For a router with d network links, each such link identifier is an integer between 1 and d. For instance, the client having address 101a has a corresponding link identifier of 3 (see **Fig. 2**).

Each routing table is constructed via information from the network topology (e.g., the client addresses) and hence from the corresponding spanning tree (e.g., the link identifiers), in a known manner.

By way of example, **Fig. 3** depicts one embodiment of a publish/subscribe system, generally denoted 300, to employ resilient message routing in accordance with the principles of the present invention. System 300 includes a network of routers 302 connected via links 303 in an arbitrary graph topology. A number of clients connected to the periphery of this router network either publish messages or subscribe to messages. Those clients publishing messages comprise publishers 304, while those clients subscribing to messages comprise subscribers 306. The router network is responsible for routing messages from a publisher 304 to interested subscribers 306 based on matching messages to subscription predicates.

This protocol, known as content-based routing, is described in detail in the initially-incorporated co-pending patent application entitled "Routing Messages Within A Network Using The Data Content Of The Message." Briefly explained, from each router node at which a publisher is present, the system computes and stores a spanning tree to reach every other node in the network. All published messages from the publisher follow the paths in that spanning tree, with each router node performing enough matching to determine which of its child routers should receive the message.

The present invention achieves reliability of routed messages by saving messages to persistent storage within the router network and retrieving and redelivering the message whenever there is a failure in the network. Publishers and subscribers that need reliability of messages may specify a quality of service parameter, e.g., "uniform delivery". Uniform delivery is provided for ensuring delivery of a message to all active subscribers notwithstanding failure in the network, e.g., the routers, or the links. Special routers in the router network 300 are designated logging nodes or loggers 310 and support the ability to log messages to stable storage. When there is at least one subscriber needing logging, the routing algorithm ensures that messages are routed to a logger. When reliable delivery is required by at least one publisher or subscriber, the routing protocol incorporates message logging in accordance with the principles of the present invention.

Briefly explained, messages originally sent by a publisher are assigned unique ids by the first router node receiving the message. All routers (including logging nodes) receiving a published message thereafter note the message's unique message id and the source node (i.e., publisher). Routing information comprising the nodes to which the message will be routed is then computed based on message content and stored, e.g., indexed by using the message id. Routing computation always includes the nearest logging node. Messages are then forwarded to applicable neighboring routers and delivered to any subscribing clients to that node not requiring reliable delivery. Messages to the node's subscribing clients requiring reliable delivery are not delivered, but instead buffered in a client message

buffer (see **Fig. 6a**) for later delivery as explained further below.

Fig. 4 depicts one embodiment of a router node for use in accordance with the principles of the present invention. When a message arrives at router 400 it is stored into a message table 402 using the unique message identifier. Information stored includes the source node identification as well as the neighboring nodes to which the message is to be forwarded. These neighboring nodes are calculated by a content routing computation component 404 after the message has been received. Computation component 404 takes the message and based upon stored subscriptions returns a set of destinations or links upon which the message should be forwarded. Again, in one embodiment, this computation is content dependent and can be accomplished as described in the above-incorporated co-pending application. However, this component can be implemented in other ways, e.g., subject-based routing. In accordance with the principles of the present invention, router 400 also includes a reliable routing component 406 and logging acknowledge (LACK) received table 408 and LACK send table 409, which are described further below.

As noted, pursuant to this invention one or more selected routers in the network also serve as logging nodes. **Fig. 5** depicts one embodiment of a logging node or logger which comprises a router 450 and persistent or stable storage 460, such as a file system or a database. When a published message is received at any logger node, the logger node performs the following logging functions in addition to the standard routing functions described above.

1. The message is assigned a next logging number in sequence.
2. The message is written to disk 460 using its logging number.
3. A logging acknowledgment (LACK) is sent to those routers to which the message was routed, as well as back towards the source node of the message. A LACK message (**Fig. 6b**) includes the message id, the logger id, and the logging number.

When a LACK message is received by a router, it looks up the routing information for the original message, sends the LACK message down the links the original message was sent excluding the link from which the LACK message was received. In addition, the router forwards the LACK towards the source of the original message, and deletes the routing information stored in the message table for that message. If the message table was lost due to the router failing and recovering, the LACK message is forwarded along all links on the spanning tree with the logger at the root, excluding the link from which the LACK message was received. Also, if the original message corresponding to the LACK was buffered waiting to be delivered to one or more subscribing clients requiring uniform delivery, the message will now be delivered.

In support of automatic recovery from failure, each router node also stores (pursuant to the present invention) the following information while processing LACK messages.

- For each logger node from which a LACK has been received, the latest logging number that has been received for that logger.
- For each link on which a LACK has been forwarded, the latest logging number that has been sent.

This information is stored into the LACK tables 408, 409 (**Fig. 4**) for the router. Since pursuant to this invention FIFO links are assumed between a logger and any given router, receiving (or sending) a LACK for a logging number N implies that all relevant LACKs for logging numbers less than N have also been received (or sent). Note that non-first-in first-out (non-FIFO) links are made FIFO for the present invention, for example, using standard techniques such as the TCP protocol.

As implied in the above description, three processing phases are implemented in one embodiment in accordance with the principles of the present invention. In a first phase, messages are routed from a publisher to all nodes in the network, but not necessarily delivered to all subscribers, depending upon whether uniform delivery has been specified by the publisher or the subscribers. In a second phase, messages are delivered from the routing nodes based upon an acknowledgment received from the logger node (i.e., the LACK message). This phase two processing and use of the loggers ensures uniform delivery of messages requiring reliable routing. A third phase is entered whenever a node or link fails within the network. This phase involves a request for and a replaying of messages that were lost due to the failure.

Fig. 7 depicts one embodiment of phase one message routing processing in accordance with the present invention. This processing is implemented at each node of the network. At a given node, a published message is received 700 and recorded into the node's message table 402 (**Fig. 4**) using the message id and source node information provided with the message 710. The node's content routing computation 404 (**Fig. 4**) next determines neighboring node routing information which is also stored into the message table 720. Again, this computation is based on data content of the message when employing a routing system such as described in the initially-incorporated patent application.

The message is next forwarded to applicable neighboring nodes 730 as determined by the content routing computation, and delivered to subscribing clients not requiring uniform delivery 740. The subscribing clients comprise clients coupled to the particular router node processing the published message. For example, certain subscriber clients to a particular router node may require uniform delivery, while others may not. Thus, the process allows those clients not requiring uniform delivery to immediately receive the published message. Again, as used herein, "uniform delivery" comprises a system guarantee that subscribers who have requested uniform delivery receive the message only if all other subscribers requesting uniform delivery also receive the message. Otherwise, the message is not forwarded to any subscriber requesting the uniform delivery. Either every subscriber receives the published message or no subscriber of uniform delivery receives the published message.

In order to deliver messages to clients requesting uniform delivery, the router node proceeds to buffer the message for later delivery to those clients 750. If the router node is a logger 760, the message is assigned a logging number and written to disk 770. A logging acknowledgment (LACK) is then sent to the applicable neighboring router nodes (computed in step 720), as well as back to the source node providing the message now held in the message table 780. This completes 790 this embodiment of the message routing phase.

Fig. 8 depicts an approach for processing LACK messages in accordance with the principles of the present invention. Each router node in the network implements this processing upon receipt of a LACK message 800. First, the routing information for the original message is determined from the node's message table 810. The LACK message is then sent to those nodes where the original message was sent minus the neighboring node from which the LACK message was received and plus the original message's source node 820. The routing information corresponding to this message id is deleted from the node's message table 830, and, if the message was buffered in the client message buffer table (such as depicted in **Fig. 6a**) for delivery to one or more clients, then the message is now delivered.

Essentially, the process comprises determining whether there is a message id in the uniform delivery table corresponding to the id of the LACK message being processed. If so, then the message is retrieved from the uniform delivery table and delivered to the one or more clients of the node requesting uniform delivery. Thus, uniform

delivery means that the message is buffered in the client message buffer of the node and then delivered once the logging acknowledgment signal is received back from a logging node of the network.

In a third phase, a routing system in accordance with the principles of the present invention enters a failure recovery process whenever a router detects failure of, for example, its parent node, or the link to its parent node. The detecting child router notifies a configuration manager, which we assume tracks the configuration of the router network. The configuration manager replies with a replacement parent node. This new parent has the same parent and children as the failed node. Implementation of this monitoring for node failure and replacement of a parent node are within the capabilities of one of ordinary skill in the art.

Note that a configuration manager is a "network system service" that is responsible for maintaining information about the structure, function, and status of a network system. This includes information such as the network topology, nodes with certain properties (such as loggers), quality of service offered nodes in the network, etc. A system's configuration may be statically fixed or dynamically changing. A configuration manager may be centralized or distributed. Network system management products such as IBM's Tivoli TME-10 contain configuration management components. From the perspective of the current invention, the system uses the topology information in the configuration manager when nodes or links fail and substitutes nodes to re-establish these failed components.

Referring to **Fig. 9**, in one embodiment, the failure recovery phase of the present invention begins with detection of failure of the parent node, or a link to the parent node 900. The configuration manager or system replaces the failed node 910, and child nodes of the failed node notify the new parent of the latest logging numbers received from each logger of the system 920. The new parent collects these responses into a sequence and forwards a request for retransmission (RTREQ) for this range of messages to its parent node 930. Upon receiving an RTREQ message, each router narrows the set of requested messages based on its logging sequence vector for the child link from which the RTREQ was received, after which the RTREQ message is forwarded on to the next node 940.

Upon receiving an RTREQ message at a logger node, the logger node narrows the set of requested messages and forwards the message onto the rest of the tree, while also finding all messages in the narrowed set of messages that it has logged and retransmits them to the new parent 950. When the new parent receives re-transmitted messages it requested from a logging router, it matches the messages that fall in the persistence range and delivers them to the appropriate children nodes 960. After all such messages have been delivered, the new parent node begins to forward new messages to its children, thereby preserving the FIFO property of the protocol.

Optionally, each router node may cache a set of recently sent messages. When used, the cache may reduce the number of messages requested in an RTREQ message or may eliminate the need for an RTREQ message.

To summarize, using the above protocol a message is guaranteed to be received by all subscribers requiring the "uniform delivery" quality of service, or none at all will receive the message. However, if a publisher crashes and recovers and as a result republishes a message, the message may be delivered multiple times to subscribers. The present invention avoids this problem using the following augmentation to the uniform delivery protocol. The augmentation is referred to as "at most once delivery" of messages, and one embodiment of the process is depicted in **Fig. 10.**

Clients desiring duplicate detection include a client-specific unique identifier with each message before forwarding the message up towards the logger node as in the original protocol 1000. Upon receipt of a data message, the logging engine checks to see if the node has previously seen the client-specific unique identifier 1010. If it has, the node logs the fact that it will abort this message, and transmits a logging abort (LABORT) message through the tree instead of a LACK message 1020. Each LABORT message still has a logging sequence number so that, in case of failures, routing nodes will receive possibly lost LABORT messages and, therefore, eventually remove routing information saved for the duplicate message. Upon receipt of a LABORT message at a node, the routing information for the original message is determined from the node's message table 1030. The LABORT message is then sent to those nodes where the original message was sent minus the neighboring node from which the LABORT message was received plus the original message's source node 1040. The routing information corresponding to this message id is deleted from the node's

message table 1050, and, if the message was buffered in the client message buffer table for delivery to one or more clients, then the message is deleted 1060.

Using the above protocol, a message is guaranteed to be received by all subscribers requiring the "uniform delivery" quality of service if the message reaches a logger node. It may also happen that a message is lost before it reaches a logger, due to failure of a router or link in route to a logger. In this case, the message is not at all delivered to uniform delivery subscribers. It is also desirable for publishers to become aware of the fact that a message has not reached a logger node, and if so, republish the message, thereby guaranteeing the delivery of the message "at least once" to every subscriber. The present invention handles this as depicted in **Fig. 11.**

Basically, a publisher can detect whether a message has been delivered to "uniform delivery" subscribers by subscribing to its own messages under the "uniform delivery" quality of service 1100. If the publisher receives the message, then it has been delivered to the uniform subscribers 1110. However, if the publisher does not receive the message within a certain time out period, it can query a logger node whether the published message has been logged 1120. If the reply to this query is received by the publisher before the original published message, it can be inferred, due to the FIFO property of links, that the original published message has been lost before it reached the logger node 1130. In this case, the message can be republished 1140.

As a further variation, "exactly once" delivery of messages can be ensured by using in combination the protocols for at most once delivery and at least once delivery of messages. A message is delivered exactly once to each subscriber in the face of: (1) publisher failures and recoveries; (2) node and link failures before a message gets to a logger; and (3) node and link failures after a message gets to a logger and before delivery to subscribers. These are respectively accomplished as follows: (1) a publisher includes a unique id in a message with the help of which the system detects duplicate publications of a message and removes the duplicate, as described above; (2) the publisher makes sure that a message reaches a logger by subscribing to its own messages as well as by querying the logger for lost messages, as described above; and (3) detecting node and link failures, re-establishing these nodes and links, and replaying lost messages from one or more loggers, as described above.

To summarize, those skilled in the art will note from the above description that in a publish/subscribe system supporting content-based subscription, a method to route messages based on content has been provided in a manner that is resilient to router or link failure. The publish/subscribe system may be realized as a router network connected in an arbitrary graph topology, except that the links from a logger node to any router are assumed to be FIFO. Further, a method to deliver a message to either all subscribers requesting a "uniform delivery" quality of service, or to none of them, is described above. Variations on this "uniform delivery" quality of service outlined above include an "at most once" quality of service, an "at least

once" quality of service, and an "exactly once" quality of service.

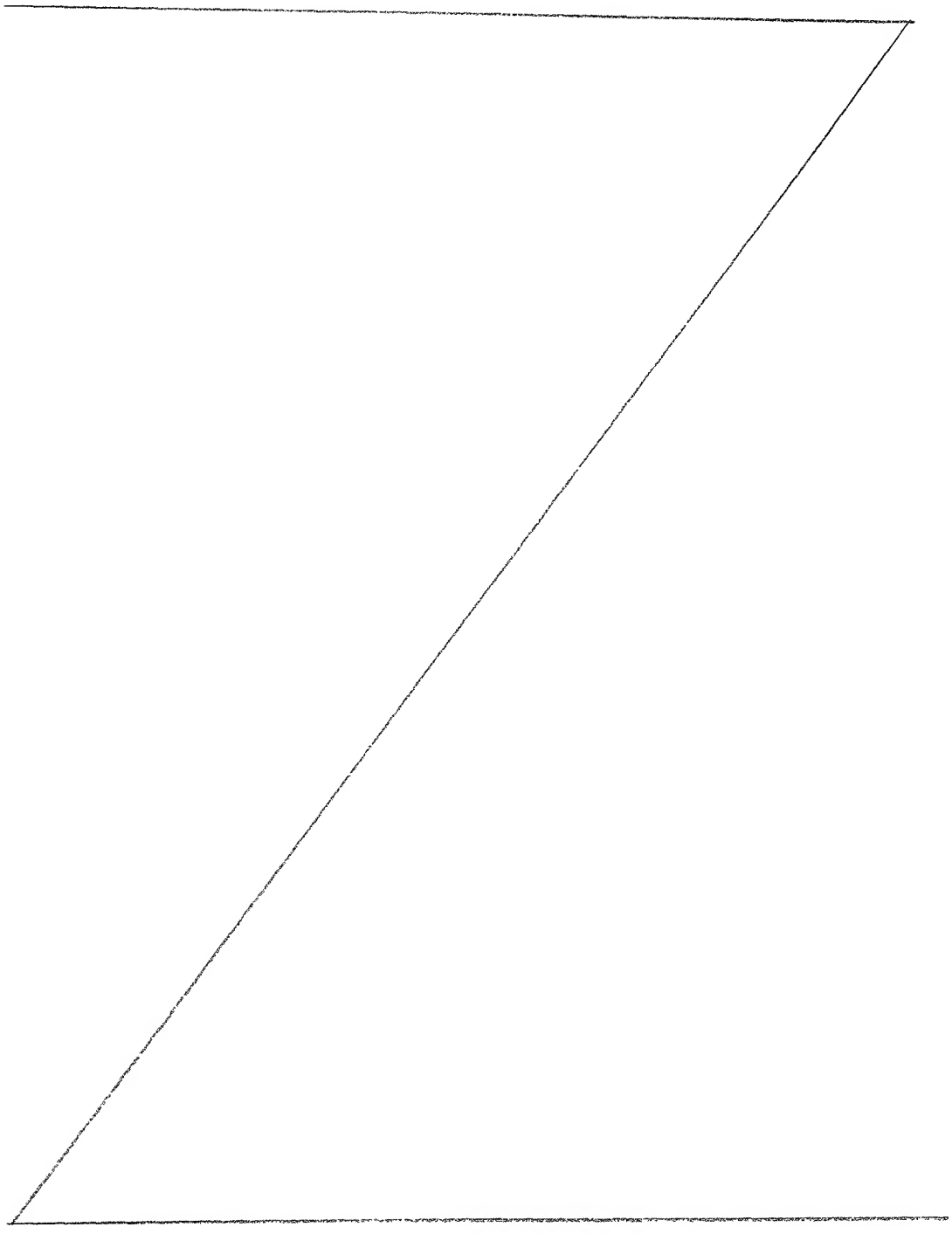
The present invention can be included, for example, in an article of manufacture (e.g., one or more computer program products) having, for instance, computer usable media. This media has embodied therein, for instance, computer readable program code means for providing and facilitating the capabilities of the present invention. The articles of manufacture can be included as part of the computer system or sold separately.

Additionally, at least one program storage device readable by machine, tangibly embodying at least one program of instructions executable by the machine, to perform the capabilities of the present invention, can be provided.

The flow diagrams depicted herein are provided by way of example. There may be variations to these diagrams or the steps (or operations) described herein without departing from the spirit of the invention. For instance, in certain cases, the steps may be performed in differing order, or steps may be added, deleted or modified. All of these variations are considered to comprise part of the present invention as recited in the appended claims.

While the invention has been described in detail herein in accordance with certain preferred embodiments thereof, many modifications and changes therein may be effected by those skilled in the art. Accordingly, it is intended by the appended claims to cover all such modifications and

changes as fall within the true spirit and scope of the invention.



Claims

1. A method for routing messages within a network, said method comprising:

receiving a message; and

routing said message to one or more clients of said network, said routing being based on data content of said message irrespective of any destination information that may be within said message, and being resilient to router or link failure within said network.

2. The method of claim 1, wherein said network comprises a publish/subscribe system supporting content-based subscription, said one or more clients comprise subscribers, and wherein said routing comprises delivering said message to all subscribers requesting a uniform delivery quality of service or if unable to deliver said message to all of said subscribers requesting uniform delivery, delivering said message to none of said subscribers requesting uniform delivery.

3. The method of claim 2, wherein said delivering said message to all subscribers requesting uniform delivery comprises delivering said message to all subscribers requesting uniform delivery notwithstanding failure at one or more routers or links of said network, said delivering comprising storing said message to persistent storage at a logging node of said network prior to providing said message to said subscribers requesting uniform delivery.

4. The method of claim 1, wherein said routing comprises logging said message at at least one logging node within said network before delivering said message to said one or more clients of said network, said logging comprising storing said message in persistent storage.

5. The method of claim 4, further comprising subsequent to said logging of said message, sending a logging acknowledgment to at least one router of said network routing said message, and upon receipt of said logging acknowledgment at said at least one router, delivering said message to a client thereof, said client requiring uniform delivery and comprising one client of said one or more clients.

6. The method of claim 5, further comprising buffering said message at said at least one router of said network routing said message, said buffering occurring prior to said storing of said message at said persistent storage and when passing said message through said at least one router to said at least one logging node.

7. The method of claim 5, wherein said network comprises a plurality of routers coupled together, one of said routers comprising said logging node having said persistent storage associated therewith, said logging comprising employing said logging node having said persistent storage associated therewith to store said message and to thereafter send said logging acknowledgment back to each router of said network responsible for routing said message.

8. The method of claim 1, wherein said network comprises a spanning tree and wherein said method further comprises providing a logging node within said spanning tree for logging said message to persistent storage during routing of said message to said one or more clients of said network.

9. The method of claim 8, wherein said routing comprises employing said logging of said message to persistent storage to ensure a uniform delivery quality of service of said message to said one or more clients of said network notwithstanding failure of one or more routers or links within said network.

10. The method of claim 1, wherein said network comprises a spanning tree having a plurality of routers, said method further comprising detecting failure of a router within said tree before completing routing of said message to said one or more clients of said network, reconfiguring said tree to replace said failed router with a new router, and automatically generating a request for retransmission of said message.

11. The method of claim 10, further comprising prior to said detecting of said failure, logging said message within persistent storage of said network and issuing a logging acknowledgment confirming storage of said message to at least one router of said tree through which said message is routed to said one or more clients.

12. The method of claim 10, wherein said automatically generating said request for retransmission of said message occurs if said new router detects from one or more of its child routers a logging number associated with said message, said logging number having been received in said logging acknowledgment confirming storage of said message.

13. The method of claim 1, wherein said routing further comprises determining within said network whether said message comprises a duplicate message to said one or more clients of said network, and if so, aborting said duplicate message such that said message is delivered to said one or more clients at most once.

14. The method of claim 1, further comprising automatically informing a sender of said message when the message has been lost within the network to allow the sender to retransmit said message for routing to said one or more clients of said network so that said message is delivered at least once to said one or more clients.

15. The method of claim 1, wherein said routing comprises logging said message at at least one logging node within said network before delivering said message to said one or more clients of said network, said logging comprising storing said message into persistent storage, and wherein said method further comprises subsequent to said logging of said message, sending a logging acknowledgment to at least one router of said network routing said message, and upon receipt of said logging acknowledgment at said at least one router of said network routing said message, looking up routing information for said message from a message table maintained at said at least one router, then sending said logging acknowledgment across said network using said looked up routing information, and thereafter deleting said routing information from said message table.

16. A method for routing messages within a routing network, said method comprising:

receiving a message;

logging the message to persistent storage within the routing network; and

after said logging, delivering said message to one or more clients of said network, wherein said logging to persistent storage prior to delivery of said message to said one or more clients of said network provides resiliency to said routing notwithstanding router or link failure within said network.

17. The method of claim 16, wherein said logging comprises storing said message in said persistent storage at a logging node within said routing network before said delivery of said message to said one or more clients of said network.

18. The method of claim 17, further comprising sending a logging acknowledgment to at least one router of said network routing said message after said logging of said message to said persistent storage, and upon receipt of said logging acknowledgment at said at least one router of said network routing said message, delivering said message to a client thereof, said client thereof requiring uniform delivery and comprising one client of said one or more clients.

19. The method of claim 16, wherein said network comprises a spanning tree and wherein said method further comprises providing a logging node within said spanning tree for said logging of said message to persistent storage during routing of said message to said one or more clients of said network.

20. The method of claim 19, wherein said routing comprises employing said logging of said message to persistent storage to ensure a uniform delivery quality of service of said message to said one or more clients of said network notwithstanding failure of one or more routers or links within said network.

21. The method of claim 16, wherein said routing network comprises a spanning tree having a plurality of routers, said method further comprising detecting failure of a router within said tree before completing routing of said message to said one or more clients of said network, thereafter reconfiguring said tree to replace said failed router with a new router, and automatically generating a request for retransmission of said message from said persistent storage.

22. The method of claim 16, further comprising determining within said routing network whether said message comprises a duplicate message to said one or more clients of said network, and if so, aborting said duplicate message such that said message is delivered to said one or more clients at most once.

23. The method of claim 16, further comprising automatically informing a sender of said message when said message has been lost within the routing network to allow the sender to retransmit the message for routing to said one or more clients of said network so that said message is delivered at least once to said one or more clients.

24. A system of routing messages within a network, said system comprising:

means for receiving a message; and

means for routing said message to one or more clients of said network, said routing being based on data content of said message irrespective of any destination information that may be within said message, and wherein said means for routing is resilient to router or link failure within said network.

25. The system of claim 24, wherein said network comprises a publish/subscribe system supporting content-based subscription, and wherein said one or more clients comprise subscribers, with said message being received from a publisher.

26. The system of claim 25, wherein said means for routing comprises means for delivering said message to all subscribers requesting a uniform delivery quality of service or if unable to deliver said message to all of said subscribers requesting uniform delivery, for delivering said message to none of said subscribers requesting uniform delivery.

27. The system of claim 26, wherein said means for routing delivers said message to said subscribers requesting uniform deliver notwithstanding failure at one or more routers or links of said network, and wherein said system further comprises means for logging said message to persistent storage prior to delivery thereof to said subscribers requesting uniform delivery.

28. The system of claim 24, wherein said means for routing comprises means for logging said message to at least one logging node within said network before delivering said message to said one or more clients of said network, said means for logging comprising means for storing said message in persistent storage.

29. The system of claim 28, further comprising means for sending a logging acknowledgment to at least one router of said network routing said message after said means for logging stores said message in persistent storage, and wherein said system further comprises, at said at least one router of said network routing said message, means for delivering said message to a client thereof upon receipt of said logging acknowledgment, said client requiring uniform delivery and comprising one client of said one or more clients.

30. The system of claim 29, further comprising means for buffering said message at said at least one router of said network routing said message, said buffering occurring prior to said storing of said message at said persistent storage by said means for logging.

31. The system of claim 29, wherein said network comprises a plurality of routers coupled together, one of said routers comprising said logging node having said persistent storage associated therewith, and wherein said means for logging comprises means for employing said logging node having said persistent storage associated therewith to store said message and to thereafter send said logging acknowledgment back to each router of said network responsible for routing said message.

32. The system of claim 24, wherein said network comprises a spanning tree and wherein said system further comprises a logger node within said spanning tree for logging said message to persistent storage during routing of said message to said one or more clients of said network.

33. The system of claim 24, wherein said means for routing comprises means for employing said logger node to log said message to persistent storage to ensure a uniform delivery quality of service of said message to said one or more clients of said network notwithstanding failure of one or more routers or links within said network.

34. The system of claim 24, wherein said network comprises a spanning tree having a plurality of routers, and further comprising means for detecting failure of a router within said tree before completing routing of said message to said one or more clients of said network, and means for reconfiguring said tree to replace said failed router with a new router, and means for automatically generating a request for retransmission of said message.

35. The system of claim 34, further comprising means for logging said message within persistent storage of said network and for issuing a logging acknowledgment confirming storage of said message to at least one router of said tree through which said message is routed to said one or more clients.

36. The system of claim 35, wherein said means for automatically generating a request for retransmission of said message comprises means for detecting a logging number associated with said message stored at one or more child routers of said new router.

37. The system of claim 24, wherein said means for routing further comprises means for determining within said network whether said message comprises a duplicate message to said one or more clients of said network, and if so, for aborting said duplicate message such that said message is delivered to said one or more clients at most once.

38. The system of claim 24, further comprising means for automatically informing a sender of said message when said message has been lost within said network to allow the sender to retransmit said message for routing to said one or more clients of said network so that said message is delivered at least once to said one or more clients.

39. A system of routing messages within a routing network, said system comprising:

means for receiving a message;

means for logging the message to persistent storage within the routing network; and

means for delivering said message to one or more clients of said network after said logging of said message to persistent storage, wherein said logging to persistent storage prior to delivery of said message to said one or more clients of said network provides resiliency to said routing notwithstanding router or link failure within said network.

40. The system of claim 39, wherein said means for logging comprises means for storing said message in said persistent storage at a logging node within said routing network before said delivery of said message to said one or more clients of said network.

41. The system of claim 40, further comprising means for sending a logging acknowledgment to at least one router of said network routing said message after said logging of said message to said persistent storage, and means for thereafter delivering said message to a client of said at least one router of said network routing said message, said client requiring uniform delivery and comprising one client of said one or more clients.

42. The system of claim 39, wherein said routing network comprises a spanning tree, and said means for logging comprises a logging node within said spanning tree for logging said message to persistent storage during routing of said message to said one or more clients of said network.

43. The system of claim 42, wherein said means for routing comprises means for employing said logging of said message to persistent storage to ensure a uniform delivery quality of service of said message to said one or more clients of said network notwithstanding failure of one or more routers or links within said network.

44. The system of claim 39, wherein said routing network comprises a spanning tree having a plurality of routers, and wherein said system further comprises means for detecting failure of a router within said tree before completing routing of said message to said one or more clients of said network, and for thereafter configuring said tree to replace said failed router with a new router, and for automatically generating a request for retransmission of said message from said persistent storage.

45. The system of claim 39, further comprising means for determining within said routing network whether said message comprises a duplicate message to said one or more clients of said network, and if so, for aborting said duplicate message such that said message is delivered to said one or more clients at most once.

46. The system of claim 39, further comprising means for automatically informing a sender of said message when said message has been lost within said routing network to allow the sender to retransmit the message for routing to said one or more clients of said network so that said message is delivered at least once to said one or more clients.

47. A system for routing messages comprising:

a network adapted to receive and log a message to persistent storage; and

said network comprising one or more routers adapted to route said message to one or more clients of said network, wherein said routing of said message by said one or more routers is based on data content of said message irrespective of any destination information that may be within the message, and is resilient to router or link failure within the network.

48. A system for routing messages comprising:

a network adapted to receive a message;

a logger node within said network for logging said message to persistent storage; and

said network comprising one or more routers for delivering said message to one or more clients of said network, wherein said logging of said message to persistent storage occurs prior to delivery of said message to said one or more clients of said network, thereby providing resiliency to said routing notwithstanding router or link failure within said network.

49. An article of manufacture, comprising:

at least one computer usable medium having computer readable program code means embodied therein for effecting routing of messages within a network, the computer readable program code means in the article of manufacture comprising:

computer readable program code means for causing a computer to effect receiving a message; and

computer readable program code means for causing a computer to effect routing said message to one or more clients of said network, said routing being based on data content of said message and being resilient to router or link failure within said network.

50. An article of manufacture, comprising:

at least one computer usable medium having computer readable program code means embodied therein for effecting routing of messages within a routing network, the computer readable program code means in the article of manufacture comprising:

computer readable program code means for causing a computer to effect receiving a message;

computer readable program code means for causing a computer to effect logging said message to persistent storage within the routing network; and

computer readable program code means for causing a computer to effect delivering said message to one or more clients of said network after said logging thereof, wherein said logging to persistent storage prior to delivery of said message to one or more clients of said network provides resiliency to said routing notwithstanding router or link failure within said network.

* * * * *

MESSAGE LOGGING FOR RELIABLE MULTICASTING ACROSS A ROUTING NETWORK

Abstract of the Disclosure

A technique for reliably multicasting a message within a router network is provided. At least one special router in the network has associated persistent storage for logging a message being routed to one or more clients. When a message is received at this logger node, the logger places the message in persistent storage and sends a logging acknowledgment back to those routers to which the message was originally routed, as well as back towards the source of the message. The logger acknowledgment includes the message id, the logger id, and a logging number. When the logger acknowledgment is received by a router, it looks up the routing information from the original message, and sends the acknowledgment to those neighboring routers to which the original message was sent, excluding the link from which the logger acknowledgment was received. If the original message corresponding to the logger acknowledgment was buffered at this router waiting to be delivered, the message is then delivered to its client nodes. Processes for recovering from node or link failure within the router network, and for at most once delivery of messages, at least once delivery of messages, and exactly once delivery of messages are also provided.

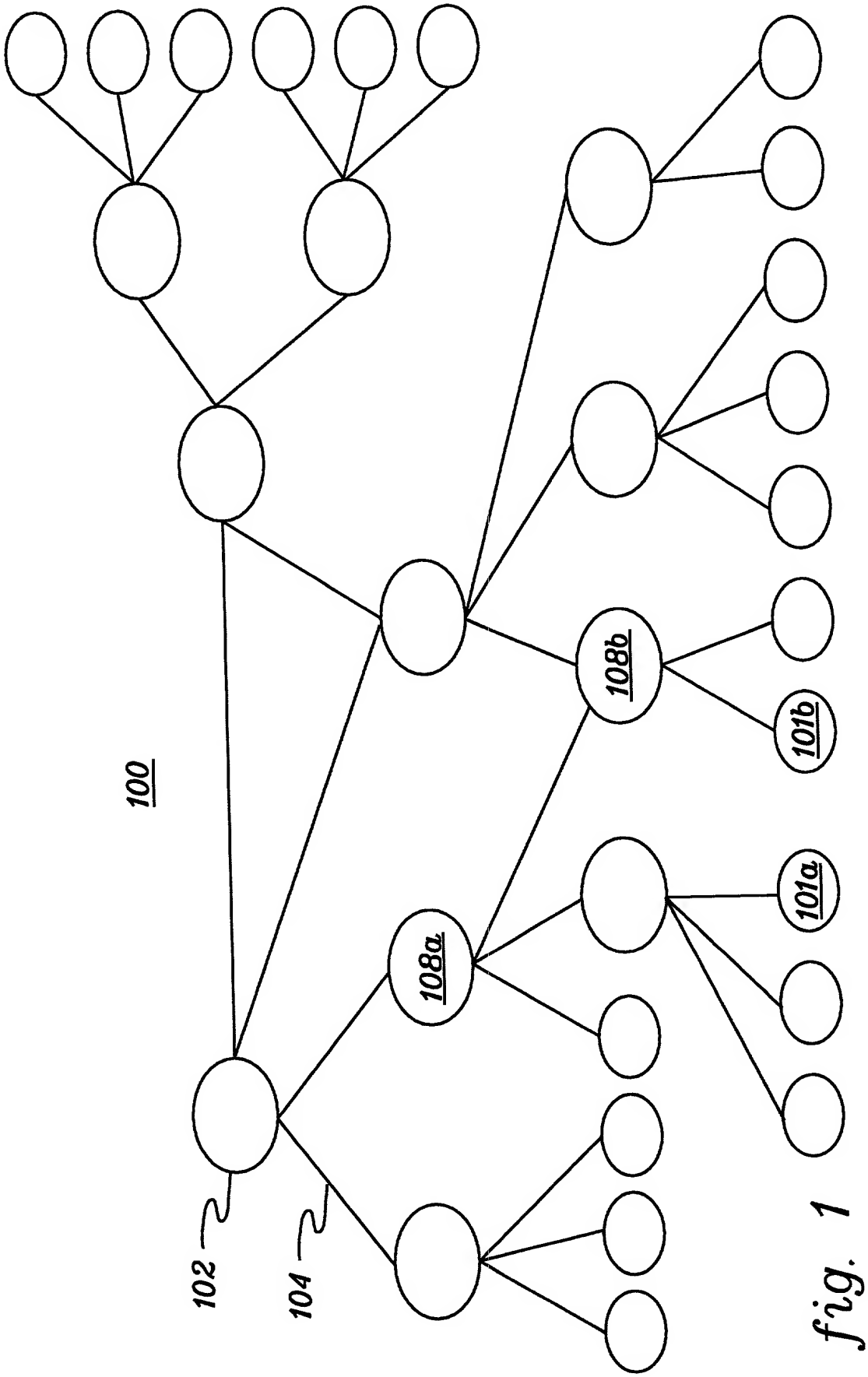


fig. 1

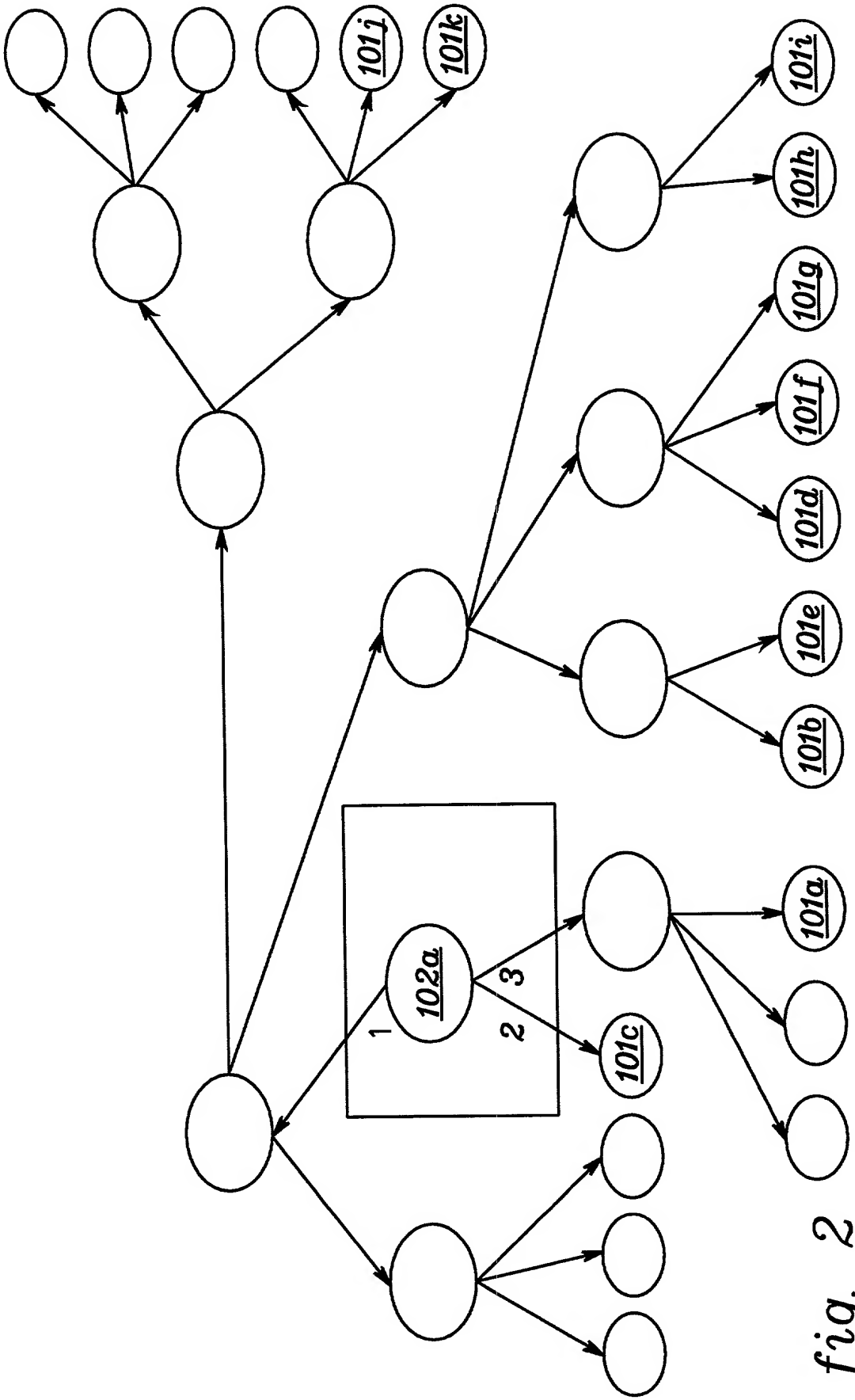

101i
101k
101i
101h
101g
101f
101d
101e
101b
101a
101c
102a
1
2
3

fig. 2

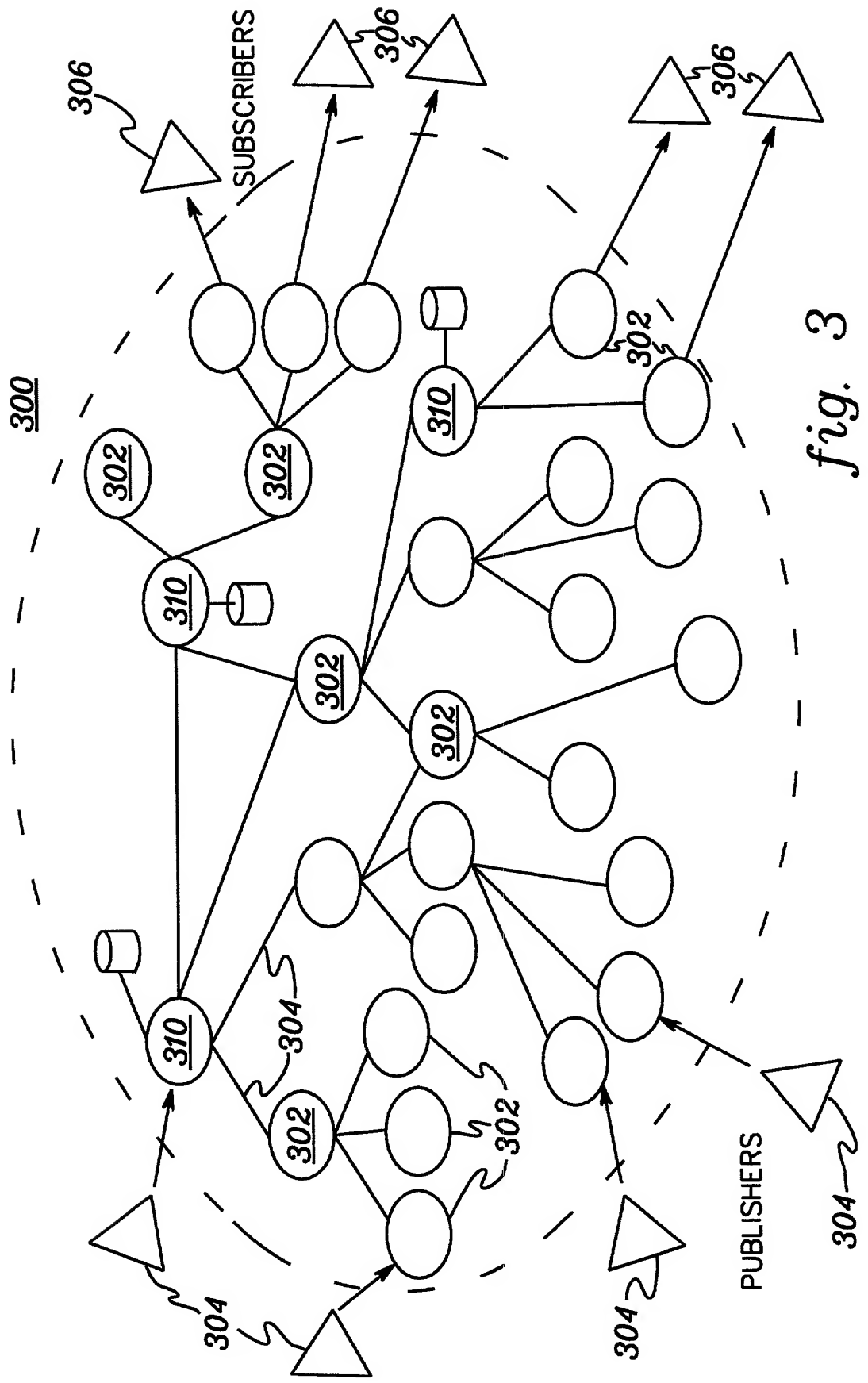

300
306
SUBSCRIBERS
306
306
302
302
310
310
302
302
310
302
304
302
302
304
304
PUBLISHERS

fig. 3

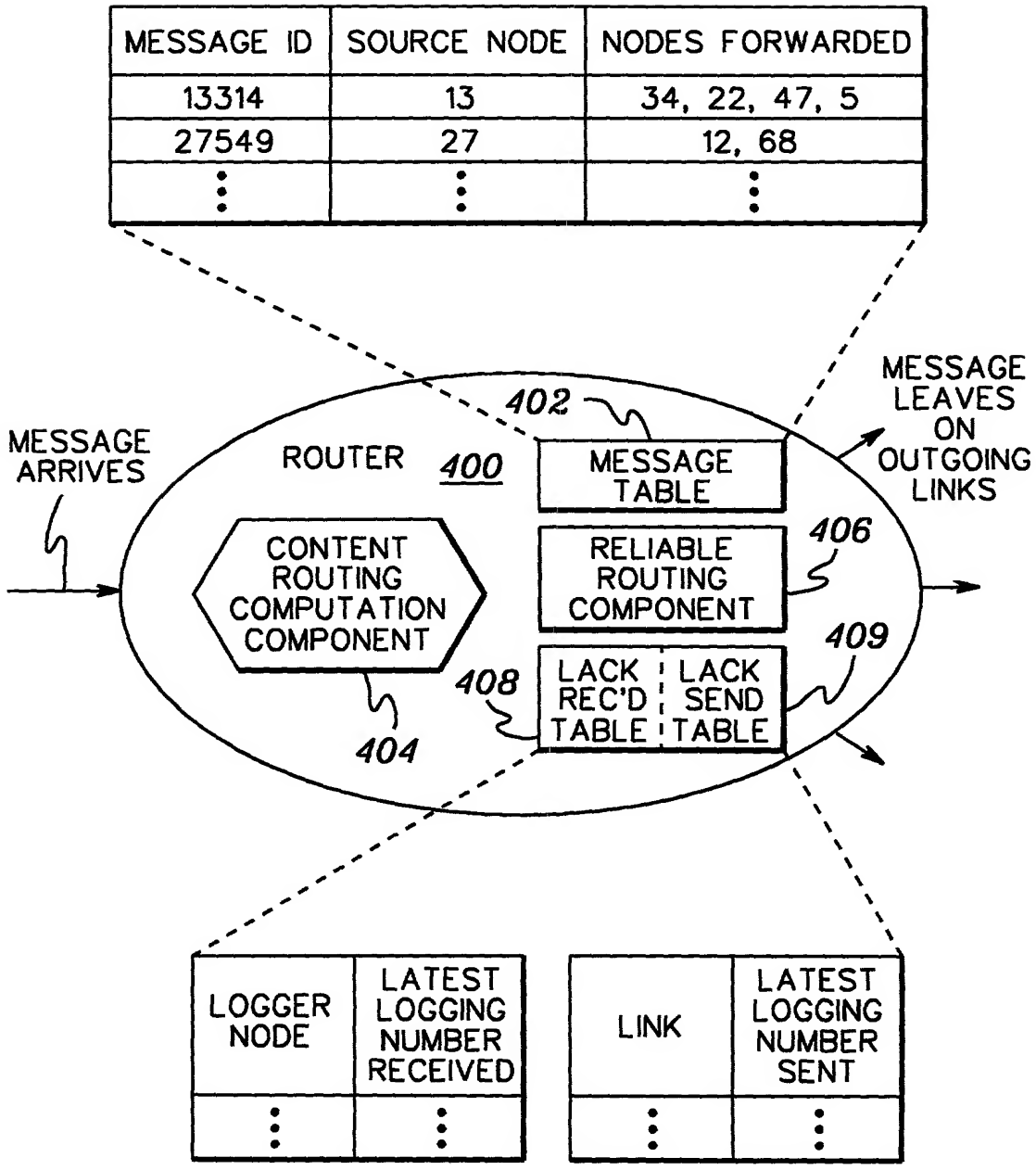

MESSAGE ID
SOURCE NODE
NODES FORWARDED
13314
13
34, 22, 47, 5
27549
27
12, 68
MESSAGE ARRIVES
ROUTER 400
402
MESSAGE TABLE
MESSAGE LEAVES ON OUTGOING LINKS
CONTENT ROUTING COMPUTATION COMPONENT
RELIABLE ROUTING COMPONENT
406
409
LACK REC'D TABLE
LACK SEND TABLE
408
404
LOGGER NODE
LATEST LOGGING NUMBER RECEIVED
LINK
LATEST LOGGING NUMBER SENT

fig. 4

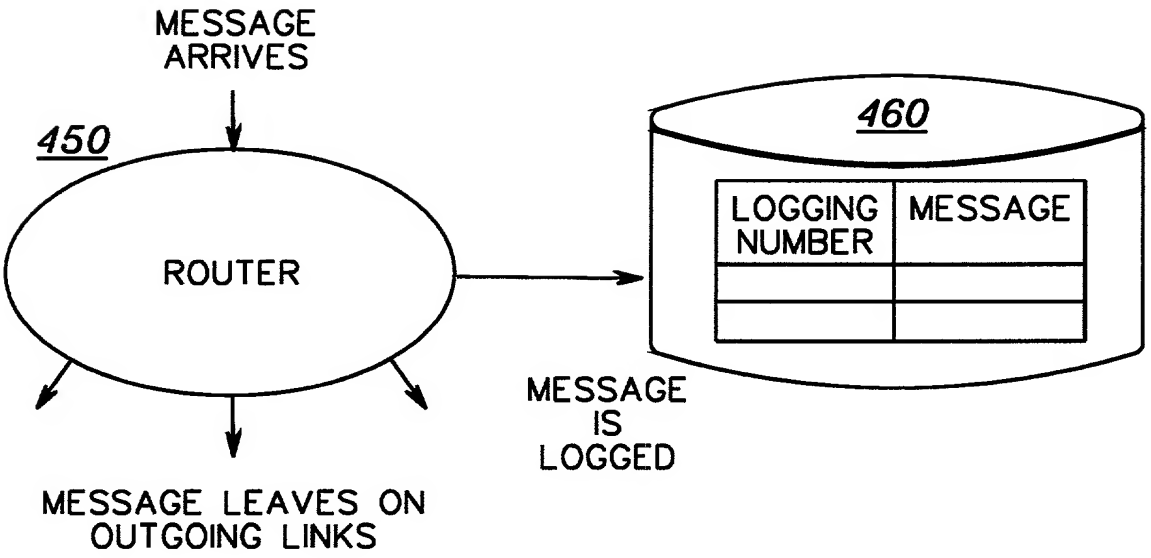


fig. 5

MESSAGE ID	MESSAGE	CLIENTS
13524	–	203, 458
3238	–	300
⋮	–	⋮

fig. 6A

MESSAGE ID	LOGGER ID	LOGGING NO.

fig. 6B

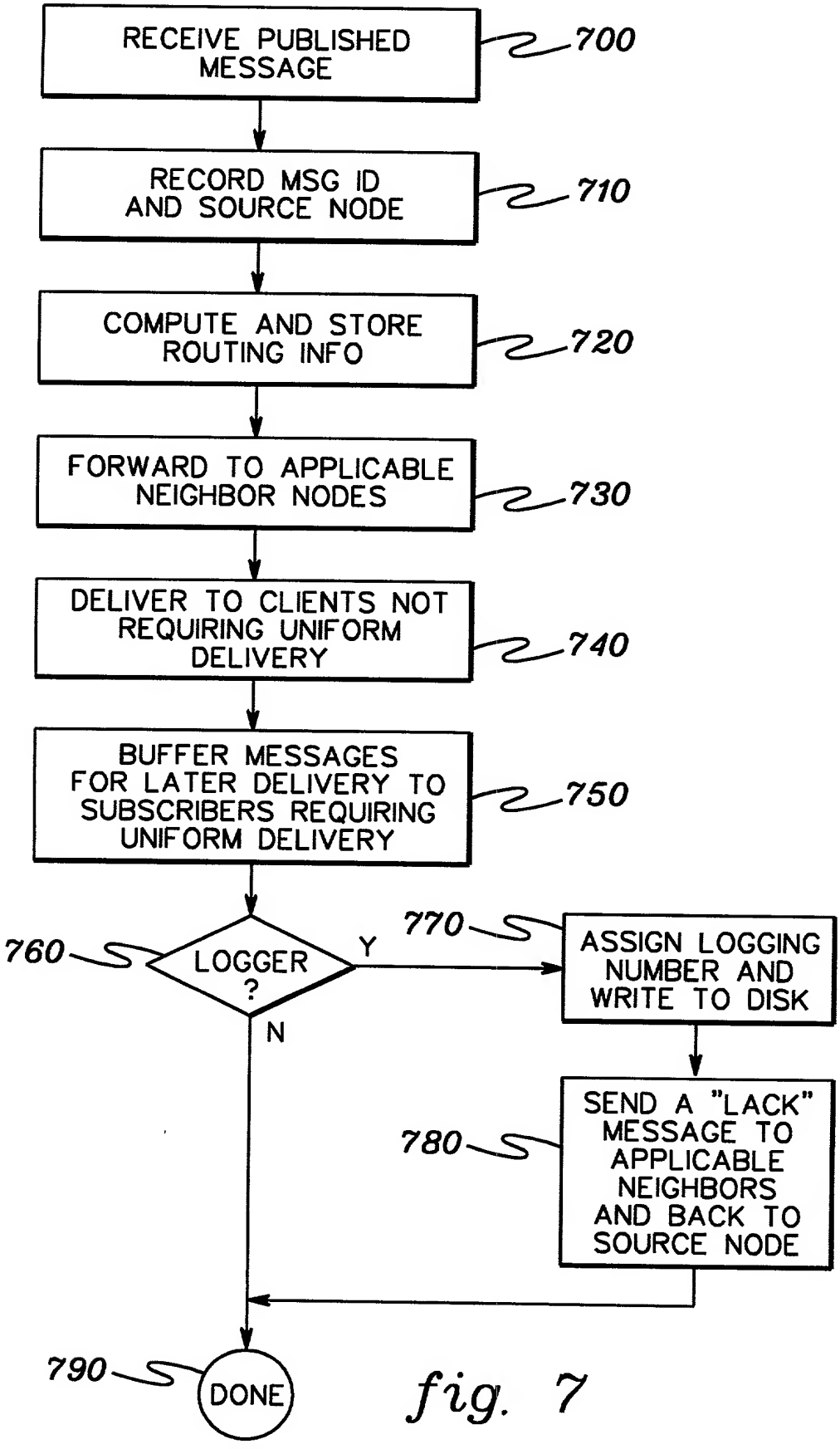

RECEIVE PUBLISHED MESSAGE
700
RECORD MSG ID AND SOURCE NODE
710
COMPUTE AND STORE ROUTING INFO
720
FORWARD TO APPLICABLE NEIGHBOR NODES
730
DELIVER TO CLIENTS NOT REQUIRING UNIFORM DELIVERY
740
BUFFER MESSAGES FOR LATER DELIVERY TO SUBSCRIBERS REQUIRING UNIFORM DELIVERY
750
760
LOGGER ?
Y
N
770
ASSIGN LOGGING NUMBER AND WRITE TO DISK
780
SEND A "LACK" MESSAGE TO APPLICABLE NEIGHBORS AND BACK TO SOURCE NODE
790
DONE

fig. 7

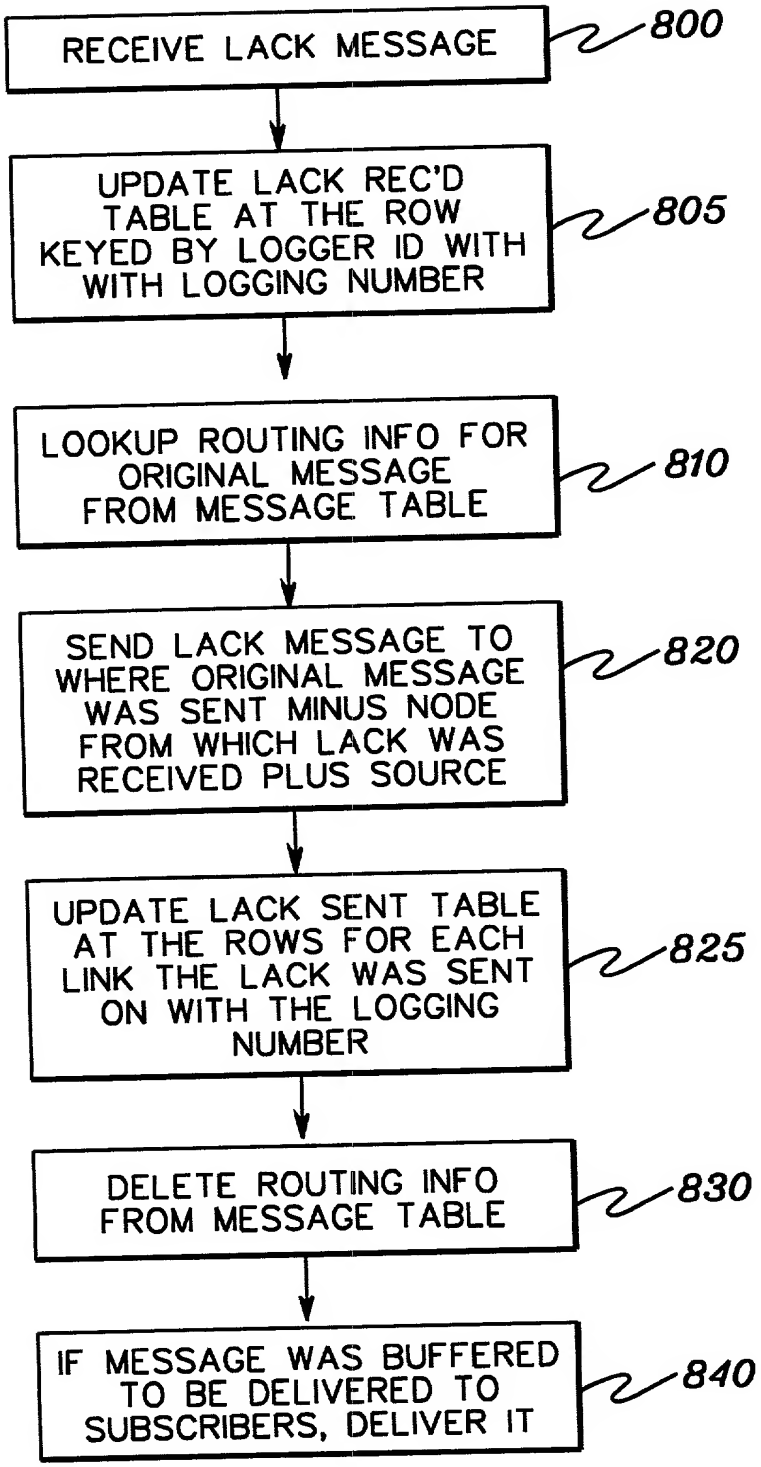

RECEIVE LACK MESSAGE
800
UPDATE LACK REC'D TABLE AT THE ROW KEYED BY LOGGER ID WITH WITH LOGGING NUMBER
805
LOOKUP ROUTING INFO FOR ORIGINAL MESSAGE FROM MESSAGE TABLE
810
SEND LACK MESSAGE TO WHERE ORIGINAL MESSAGE WAS SENT MINUS NODE FROM WHICH LACK WAS RECEIVED PLUS SOURCE
820
UPDATE LACK SENT TABLE AT THE ROWS FOR EACH LINK THE LACK WAS SENT ON WITH THE LOGGING NUMBER
825
DELETE ROUTING INFO FROM MESSAGE TABLE
830
IF MESSAGE WAS BUFFERED TO BE DELIVERED TO SUBSCRIBERS, DELIVER IT
840

fig. 8

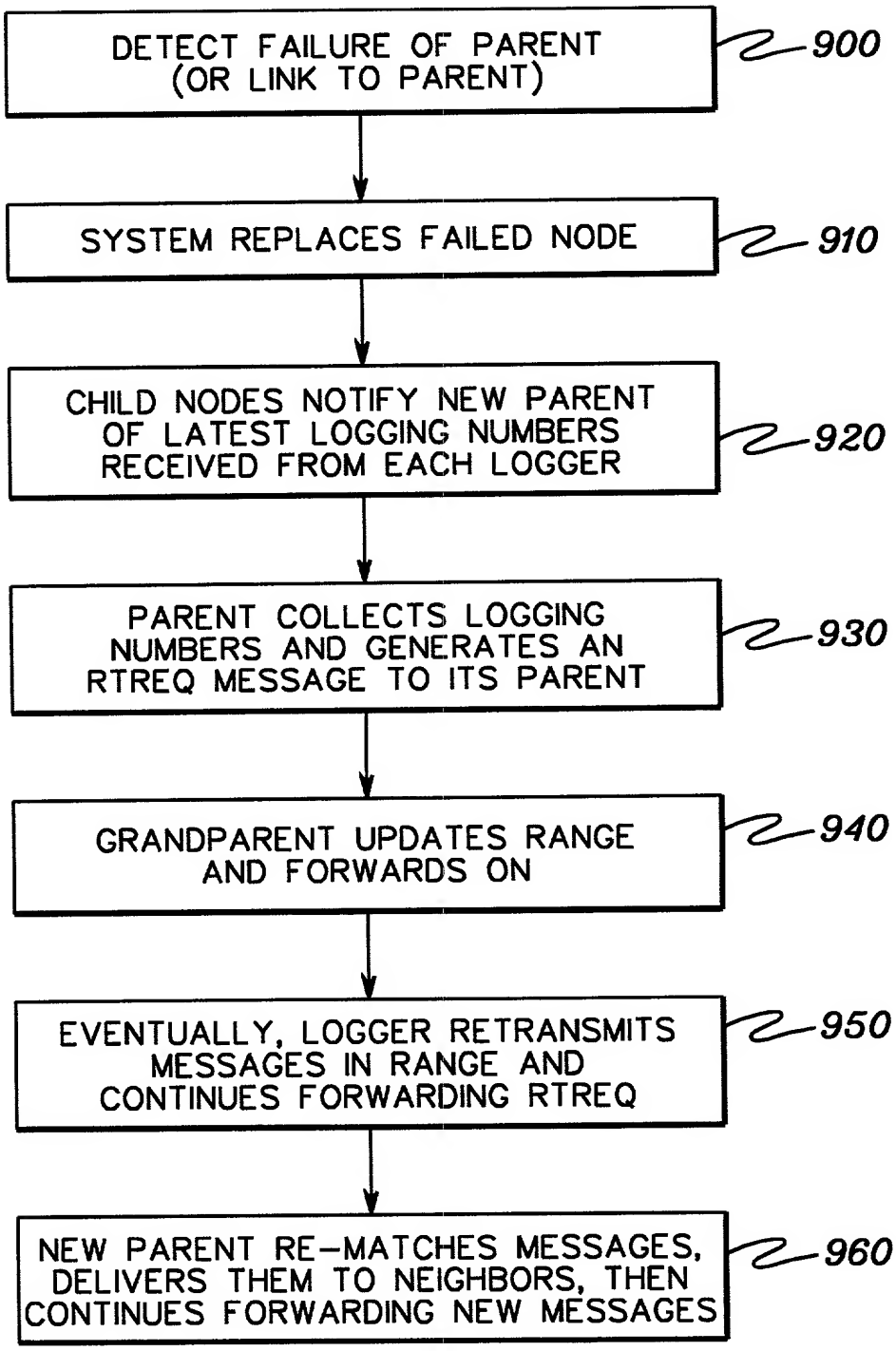

DETECT FAILURE OF PARENT (OR LINK TO PARENT)
900
SYSTEM REPLACES FAILED NODE
910
CHILD NODES NOTIFY NEW PARENT OF LATEST LOGGING NUMBERS RECEIVED FROM EACH LOGGER
920
PARENT COLLECTS LOGGING NUMBERS AND GENERATES AN RTREQ MESSAGE TO ITS PARENT
930
GRANDPARENT UPDATES RANGE AND FORWARDS ON
940
EVENTUALLY, LOGGER RETRANSMITS MESSAGES IN RANGE AND CONTINUES FORWARDING RTREQ
950
NEW PARENT RE-MATCHES MESSAGES, DELIVERS THEM TO NEIGHBORS, THEN CONTINUES FORWARDING NEW MESSAGES
960

fig. 9

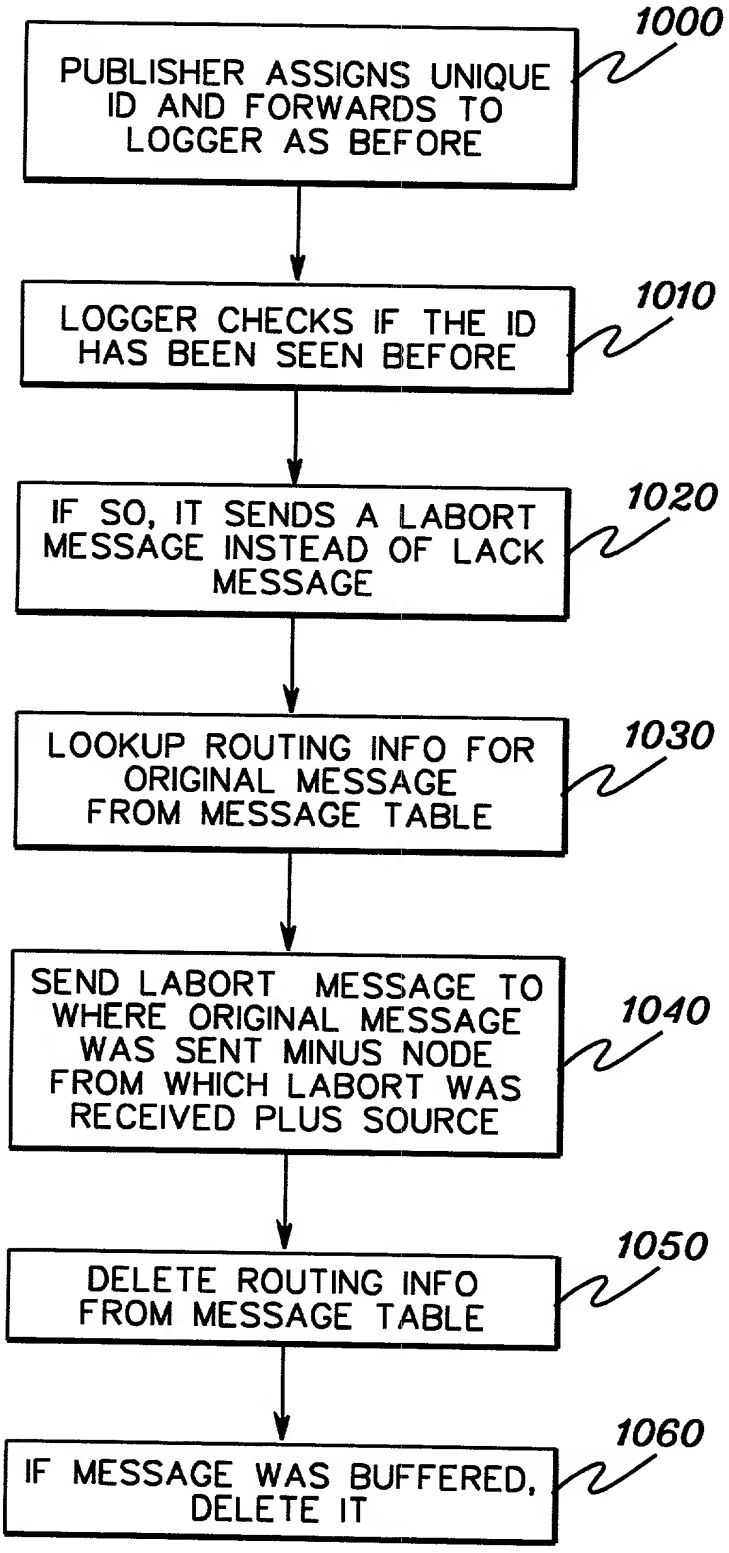

PUBLISHER ASSIGNS UNIQUE ID AND FORWARDS TO LOGGER AS BEFORE
1000
LOGGER CHECKS IF THE ID HAS BEEN SEEN BEFORE
1010
IF SO, IT SENDS A LABORT MESSAGE INSTEAD OF LACK MESSAGE
1020
LOOKUP ROUTING INFO FOR ORIGINAL MESSAGE FROM MESSAGE TABLE
1030
SEND LABORT MESSAGE TO WHERE ORIGINAL MESSAGE WAS SENT MINUS NODE FROM WHICH LABORT WAS RECEIVED PLUS SOURCE
1040
DELETE ROUTING INFO FROM MESSAGE TABLE
1050
IF MESSAGE WAS BUFFERED, DELETE IT
1060

fig. 10

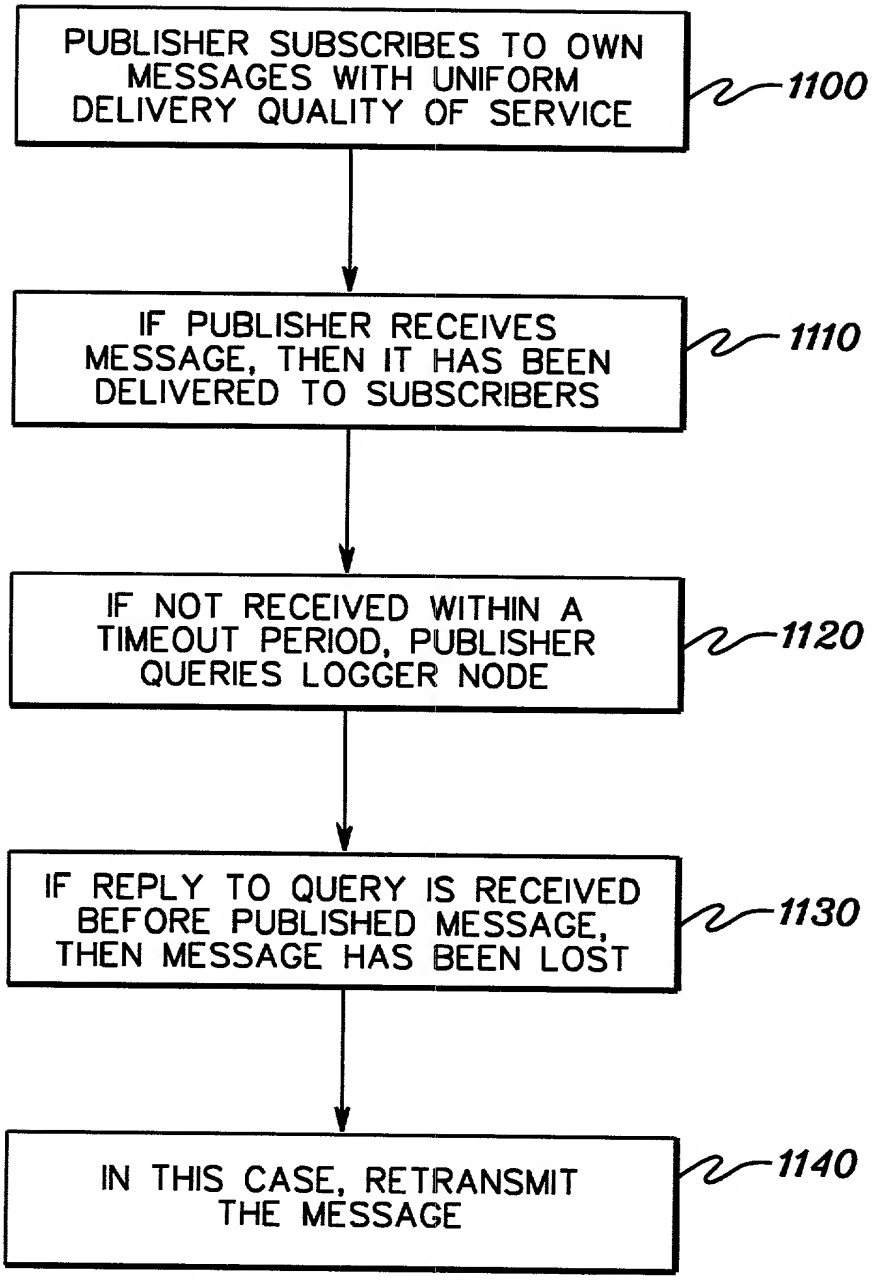

PUBLISHER SUBSCRIBES TO OWN MESSAGES WITH UNIFORM DELIVERY QUALITY OF SERVICE
1100
IF PUBLISHER RECEIVES MESSAGE, THEN IT HAS BEEN DELIVERED TO SUBSCRIBERS
1110
IF NOT RECEIVED WITHIN A TIMEOUT PERIOD, PUBLISHER QUERIES LOGGER NODE
1120
IF REPLY TO QUERY IS RECEIVED BEFORE PUBLISHED MESSAGE, THEN MESSAGE HAS BEEN LOST
1130
IN THIS CASE, RETRANSMIT THE MESSAGE
1140

fig. 11

Docket No.
YO998-525

Declaration and Power of Attorney For Patent Application

English Language Declaration

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name,

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled

MESSAGE LOGGING FOR RELIABLE MULTICASTING ACROSS A ROUTING NETWORK

the specification of which

(check one)

☒ is attached hereto.

☐ was filed on _____ as United States Application No. or PCT International Application Number _____

and was amended on _____
(if applicable)

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose to the United States Patent and Trademark Office all information known to me to be material to patentability as defined in Title 37, Code of Federal Regulations, Section 1.56.

I hereby claim foreign priority benefits under Title 35, United States Code, Section 119(a)-(d) or Section 365(b) of any foreign application(s) for patent or inventor's certificate, or Section 365(a) of any PCT International application which designated at least one country other than the United States, listed below and have also identified below, by checking the box, any foreign application for patent or inventor's certificate or PCT International application having a filing date before that of the application on which priority is claimed.

Prior Foreign Application(s)			Priority Not Claimed
_____ (Number)	_____ (Country)	_____ (Day/Month/Year Filed)	☐
_____ (Number)	_____ (Country)	_____ (Day/Month/Year Filed)	☐
_____ (Number)	_____ (Country)	_____ (Day/Month/Year Filed)	☐

I hereby claim the benefit under 35 U.S.C. Section 119(e) of any United States provisional application(s) listed below:

(Application Serial No.)	(Filing Date)
(Application Serial No.)	(Filing Date)
(Application Serial No.)	(Filing Date)

I hereby claim the benefit under 35 U. S. C. Section 120 of any United States application(s), or Section 365(c) of any PCT International application designating the United States, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT International application in the manner provided by the first paragraph of 35 U.S.C. Section 112, I acknowledge the duty to disclose to the United States Patent and Trademark Office all information known to me to be material to patentability as defined in Title 37, C. F. R., Section 1.56 which became available between the filing date of the prior application and the national or PCT International filing date of this application:

(Application Serial No.)	(Filing Date)	(Status) (patented, pending, abandoned)
(Application Serial No.)	(Filing Date)	(Status) (patented, pending, abandoned)
(Application Serial No.)	(Filing Date)	(Status) (patented, pending, abandoned)

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

POWER OF ATTORNEY: As a named inventor, I hereby appoint the following attorney(s) and/or agent(s) to prosecute this application and transact all business in the Patent and Trademark Office connected therewith. *(list name and registration number)*

Manny W. Schecter (Reg. 31,722)
Terry J. Ilardi (Reg. 29,936)
Christopher A. Hughes (Reg. 26,914)
Edward A. Pennington (Reg. 32,588)
John E. Hoel (Reg. 26,279)
Joseph C. Redmond, Jr. (Reg. 18,753)
Kevin M. Jordan (Reg. 40,277)
Stephen C. Kaufman (Reg. 29,551)
Jay P. Sbrollini (Reg. 36,266)
David M. Shofi (Reg. 39,835)
Robert M. Trepp (Reg. 25,933)
Daniel P. Morris (Reg. 32,053)
Douglas W. Cameron (Reg. 31,596)
Paul J. Otterstedt (Reg. 37,411)

Send Correspondence to: **Kevin P. Radigan, Esq.**
HESLIN & ROTHENBERG, P.C.
5 Columbia Circle
Albany, NY 12203

Direct Telephone Calls to: *(name and telephone number)*
Kevin P. Radigan, Esq. (518) 452-5600

Full name of sole or first inventor
GURUDUTH SOMASEKHARA BANAVAR

Sole or first inventor's signature: G Duth Somasekhara Banavar — Date: Mar 24, 1999

Residence
3148 Gomer Street, Yorktown Heights, NY 10598

Citizenship
India

Post Office Address
3148 Gomer Street, Yorktown Heights, NY 10598

Full name of second inventor, if any
TUSHAR DEEPAK CHANDRA

Second inventor's signature: Tushar Deepak Chandra — Date: March 25 '99

Residence
320 East 46th Street, Apt. 31H, New York, NY 10017

Citizenship
India

Post Office Address
320 East 46th Street, Apt. 31 H, New York, NY 10017

Full name of third inventor, if any
KEVAN LEE MILLER

Third inventor's signature Kevan Lee Miller Date March 24, 1999

Residence
3 Kilian Drive, Danbury, CT 06811

Citizenship
United States of America

Post Office Address
3 Kilian Drive, Danbury, CT 06811

Full name of fourth inventor, if any
ROBERT EVAN STROM

Fourth inventor's signature Robert Evan Strom Date March 24, 1999

Residence
6 Rochambeau Avenue, Ridgefield, CT 06877

Citizenship
United States of America

Post Office Address
6 Rochambeau Avenue, Ridgefield, CT 06877

Full name of fifth inventor, if any
DANIEL CHARLES STURMAN

Fifth inventor's signature Date March 24, 1999

Residence
408 Knickerbocker Road, Englewood, NJ 07631

Citizenship
United States of America

Post Office Address
408 Knickerbocker Road, Englewood, NJ 07631

Full name of sixth inventor, if any
MICHAEL JAMES WARD

Sixth inventor's signature Michael James Ward Date March 24, 1999

Residence
25 West Park Avenue, New Haven, CT 06511

Citizenship
United States of America

Post Office Address
25 West Park Avenue, New Haven, CT 06511

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicants: Banavar et al.

Serial No.: Group Art Unit:

Filed: Examiner:

Title: MESSAGE LOGGING FOR RELIABLE MULTICASTING ACROSS A ROUTING NETWORK

APPOINTMENT OF ASSOCIATE ATTORNEY

To: Assistant Commissioner for Patents
Washington, D.C. 20231

Sir:

The undersigned attorney, who is an attorney in the Combined Declaration and Power of Attorney for the above-identified application, hereby appoints Jeff Rothenberg, Esq., Reg. No. 26,429, Kevin Radigan, Esq., Reg. No. 31,789, and Blanche E. Schiller, Esq., Reg. No. 35,670, of Heslin & Rothenberg, P.C., 5 Columbia Circle, Albany, New York 12203 (telephone number 518-452-5600) as his associate attorneys to prosecute said application and to transact all business in the Patent and Trademark Office connected therewith.

Please direct all official communications to Kevin P. Radigan at the address cited above.

Respectfully submitted,

Kevin M. Jordan, Esq. - Attorney
Reg. No.: 40,277
Tel. No.: (914) 945-2120

IBM Corporation
Intellectual Property Law Dept.
T.J. Watson Research Center
P.O. Box 218
Yorktown Heights, NY 10598